



# Provincial income convergence in Vietnam: spatio-temporal dynamics and conditioning factors

Minh-Thu Thi Nguyen<sup>1,2</sup>

Received: 12 May 2023 / Accepted: 18 March 2024  
© The Author(s) 2024

## Abstract

In this article, we use a three-step procedure that combines the log  $t$  convergence test, Explanatory Spatial Data Analysis, and ordered logit regression to determine the spatio-temporal dynamics and determinants of provincial income clustering in Vietnam during the 2010–2020 period. Our findings are three-fold. First, provincial income clustering in Vietnam follows patterns of club convergence towards multiple equilibria. Seven convergence clubs encompassing 61 provinces are identified. Second, spatial autocorrelation encourages neighboring provinces to converge toward shared income equilibria. High-income clusters are observed in the Northern and Southern Key Economic Regions, while low-income clusters are concentrated in the mountainous areas of Northern Vietnam. Finally, both internal and external factors significantly affect the formation of convergence clubs. Vital internal factors include localities' initial conditions of physical capital and structural change. Meanwhile, external factors refer to spatial externalities among neighboring provinces. We highlight spatial complementarity in physical capital accumulation and spatial competition in industrial intensification among neighboring provinces.

**Keywords** Income clustering · Club convergence · Spatial dependence · Structural changes · Spatial externalities · Vietnam

**JEL Classification** O40 · O47 · R10 · R11

## 1 Introduction

Vietnam was one of the world's poorest countries after the end of the Vietnam War in 1975. In the late 1980s, the Vietnamese Government conducted an economic reform tackling a shift from a centrally planned to a market-oriented economy. Thanks to

---

✉ Minh-Thu Thi Nguyen  
nguyen.thi.minh.thu.z7@s.mail.nagoya-u.ac.jp; minhthunguyen2910@gmail.com

<sup>1</sup> Graduate School of International Development, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Japan

<sup>2</sup> Ho Chi Minh National Academy of Politics, Ha Noi, Vietnam

the reform, the country has successfully transformed into a vibrant economy with remarkable economic growth and poverty reduction in less than 40 years. The average annual growth rate for the 1990–2018 period was recorded at 8.8% (Bentzen and Tung 2021); meanwhile, the poverty headcount ratio plunged from 58.1% in 1992 to roughly 5% in 2018 (Pham and Riedel 2019). However, income inequality has gradually evolved into a primary development concern in Vietnam. The ratio between the average income of the 20% most affluent population over that of the 20% poorest population was recorded to increase significantly from about 8.0 times in 2004 (Oxfam 2017) to 10.23 times in 2019 (GSO 2020). As rising inequality will make it more challenging for the Vietnamese Government to pursue the pro-poor growth objective, addressing income disparity has been placed at the forefront of the development agenda, besides maintaining high economic growth. To tackle the issue of income inequality, it is crucial to comprehensively understand both the temporal and spatial sources of local income clustering in Vietnam. Motivated by this context, we aim to reveal the spatio-temporal dynamics and determinants of local income clustering in Vietnam during the 2010–2020 period. We propose a combined research framework of the log  $t$  convergence test, Explanatory Spatial Data Analysis (ESDA), and the ordered logit model to achieve the research objective. Based on this combined framework, our main findings are three-fold. First, local income clustering in Vietnam presents club convergence patterns during the studied period. Seven clubs encompassing 61 provinces are revealed. Provinces in Northern Vietnam show more heterogeneity in their convergence patterns than those in Central and Southern Vietnam. Next, we confirm the existence of spatial autocorrelation that encourages nearby provinces to converge to shared income equilibria, forming spatial income clusters. One low-income cluster has been found in the Northern mountainous areas. Two other high-income clusters have also been revealed, one in the North and one in the South. Finally, we emphasize the role of internal factors (i.e., physical capital accumulation and structural change) and external factors (i.e., spatial externalities among neighboring provinces) as critical determinants of local income clustering in Vietnam. Along the development process, neighboring provinces are complementary in terms of physical capital accumulation but competitive while striving for an industrial-based economic structure.

Our paper advances the literature on income inequality in Vietnam from several perspectives. First, it documents the temporal sources of local income clustering in Vietnam using the non-linear dynamic factor model. One common feature shared among recent studies of regional income convergence in Vietnam, such as Bentzen and Tung (2021) and Diem and Hart (2022), is that they are developed from the classical convergence framework to capture the convergence behavior of a representative economy. Beyond the average, we document unique convergence behaviors driven by regional heterogeneity or other possible non-linear factors. Second, this study combines the non-linear dynamic factor model and ESDA methods to elaborate on the spatio-temporal dynamics of local income clustering in Vietnam. Before us, only Vu et al (2018) applied the club convergence method to examine the temporal dynamics of clustering behaviors among Vietnamese provinces during the 1990–2011 period. We add extra values by revealing the spatial dynamics and identifying spatial clusters and outliers. Further investigation of spatial clusters and outliers enables us to elaborate

on critical local economies that should be focalized to address the persistently poor areas in Vietnam. These focalization implications are particularly beneficial when Vietnam, like many other developing countries, encounters a great dilemma of constrained resources versus abundant development demands. Finally, while analyzing the determinants of convergence clubs, we provide novel evidence of the impacts of spatial externalities on club formation. We identify the co-existence of opposite relationships among neighboring provinces during their development. Specifically, they are complementary in terms of physical capital accumulation but competitive in terms of industrial development. To our knowledge, this has not been acknowledged in previous literature on income convergence in Vietnam.

The rest of the paper is structured as follows: Sect. 2 provides background information about Vietnam's geographic structure and regional development policies to enhance the readability of result discussions. Section 3 discusses the related literature on regional income convergence in Vietnam, followed by methodology and data explanations in Sect. 4. Section 5 communicates results and discussions. The paper ends with concluding remarks and policy recommendations in Sect. 6.

## 2 Vietnam's geographical structure and regional development policies

From North to South, Vietnam's land territory encompasses five centrally controlled cities and 58 provinces at the first level of administrative governance. Due to their analogous jurisdictional characteristics, the five centrally controlled cities and 58 provinces are equated in our study. As such, it is possible to consider that 63 provinces constitute Vietnam's land territory. They are categorized into three regions: Northern, Central, and Southern Vietnam. Figure 8 provides a map of Vietnam with the country's three geographical regions being highlighted.

Northern Vietnam is featured by two distinct topographical types: the Northern Midlands and Mountains, covering 15 mountainous provinces in the far North and Northwest, and the Red River Delta, spanning ten plain provinces. Unlike the Northern Midlands and Mountains, the Red River Delta is destined with favorable natural conditions for economic development. Since the late 1980s economic reform, it has been designated as a critical national industrial hub. The Red River Delta is also among the most populous regions, with two municipalities directly under the Central Government's management: Ha Noi, the capital city, and Hai Phong, the North's largest seaport city. Central Vietnam, comprising 19 provinces, is a narrow land strip connecting Northern and Southern Vietnam. The long eastern coastal line establishes the prominent geographic feature of Central Vietnam. This feature allows the region to develop deep seaports, logistics, and marine services. Da Nang City acts as the core of the entire region. Finally, Southern Vietnam, consisting of 19 provinces, is characterized by the flat and "open to the sea" terrain in the Southeast and the vast Mekong River Delta in the far South. These geographical advantages have propelled Southern Vietnam to become the country's most dynamic economic region. The Mekong River Delta is the largest seafood and agricultural production center; meanwhile, the Southeastern part is home to major industrial economies, such as Ho Chi Minh City, Binh Duong, and Dong Nai.

Since the late 1990s, the Vietnamese Government has emphasized regional economic development policies as an essential pillar of the national development strategy. One key component of the regional economic development policies is the Key Economic Region (KER) policy. In 1997 and 1998, the Government established three KERs: the Northern KER in the North, the Central KER in the Central part, and the Southern KER in the South. In each KER, one municipality (i.e., Ha Noi in the North, Da Nang City in Central Vietnam, and Ho Chi Minh City in the South) is selected as the core city directing infrastructure and production linkages across the region. Involving a total of 20 provinces, the KER policy aims to foster solid economic dynamism within each KER while promoting positive economic spillovers to surrounding areas. The Government also envisions transforming KERs into vital industrial hubs, each focusing on specific industrial sectors based on regional advantages. To facilitate KERs, the Government issued Decision 20/2004/QD-TTg in 2004 and Decision 159/2007/QD-TTg in 2007 to establish the KER Development Coordination Agency and the coordination regulations among ministries, sectors, and localities regarding the development of KERs. Public investments also focus on developing transportation infrastructure and enhancing regional connectivity among KER economies. Following Vietnam's accession to the World Trade Organization in 2007, provinces in KERs have become prominent destinations for foreign direct investment. Table 12 provides an overview of the main economic features of three KERs in Vietnam.

### 3 Literature review

#### 3.1 Regional income convergence studies

This study is built upon the income club convergence framework pioneered by Phillips and Sul (2007, 2009). The framework suggests that regional heterogeneity and other non-linear factors may prevent economies from converging smoothly toward a universal income equilibrium. Instead, multiple equilibria defined by particular sets of attributes might exist, and economies possessing attributes corresponding to each income equilibrium will likely converge toward that equilibrium, forming a convergence club (Islam 2003).

Among income club convergence studies, a subset of research focuses on examining the varying patterns of convergence among different countries. For examples, Bartkowska and Riedl (2012); Von Lyncker and Thoennesen (2017), and Cutrini (2019) discuss convergence behaviors among regions in Europe. Although the screening period varies among the three papers, they all share the finding that European regions converge toward multiple equilibria. Multiple convergence clubs are driven by critical factors, such as regions' initial conditions of income per capita and growth fundamentals. Revisiting the role of structural characteristics, only an economic structure dominated by high added-values sectors, such as high-tech manufacturing and knowledge-intensive services, could promote convergence toward a higher income equilibrium (Bartkowska and Riedl 2012; Cutrini 2019). Other demographic and urban features, such as population growth rate (Von Lyncker and Thoennesen 2017)

and degree of urbanization (Cutrini 2019), also condition income convergence across European regions. Moving research to the context of Latin America and Asia, Martin and Vazquez (2015) and Tam (2018) shed light on income club convergence among countries within these two parts of the world, respectively. Martin and Vazquez (2015) identifies three convergence clubs among 18 Latin American countries from the 1950 to 2008 period. The heterogeneity in institutional quality among the countries largely explains their clustering behaviors to form different convergence clubs. Using the log  $t$  test, Tam (2018) explores the convergence patterns in income, consumption expenditure, and government expenditure among 19 Asian economies over the 1970–2014 period. In each of these macroeconomic aspects, it is found that the dynamic growth paths of Asian countries converge into two groups, which are well differentiated by the average value line.

Another strand of the income club convergence literature involves research exploring diverse convergence patterns within individual countries. Recent studies argue that persistent differences in human capital, provincial institution quality, and geography might also prevent local economies within a single country from smoothly converging toward a universal equilibrium. For example, Gunawan et al (2021) and Aginta (2021) document multiple income convergence clubs among Indonesian local economies at the provincial and district level, respectively. Zhang et al (2019) identifies four convergence clubs among 329 cities in China between 1990 and 2019. Ursavaş and Mendez (2022) also records multiple convergence clubs while studying income convergence among 81 provinces in Turkey during the 2012–2019 period. In these studies, although factors conditioning club membership might vary depending on the country-specific context, key determinants are usually found among initial conditions of income, human capital, physical capital, structural change, agglomeration effect, government spending, foreign investment, etc.

### 3.2 Spatial dependence in regional income convergence studies

Spatial dependence, a concept in spatial analysis, describes a scenario where the value of a variable observed in one location is influenced by the values of the same variable in nearby locations. In the field of regional income convergence analysis, spatial dependence suggests that the income level in one region might be affected by the income levels in the adjacent areas. The role of spatial dependence has been examined in several studies, such as Bartkowska and Riedl (2012); Zhang et al (2019), and Ursavaş and Mendez (2022). However, the findings regarding the impact of spatial dependence on convergence club formation are mixed among these studies. Bartkowska and Riedl (2012) is the first to test the impact of spatial dependence on establishing six convergence clubs across 206 European regions from 1990 to 2002. In this study, spatial dependence is statistically insignificant among the drivers of club formation. Unlike Bartkowska and Riedl (2012); Zhang et al (2019) finds that spatial effects, proxied by the average annual growth rate of neighboring regions, perform a critical role in the establishment of four convergence clubs across 329 Chinese prefectures during the 1990–2014 period. Ursavaş and Mendez (2022) also finds spatial factors essential for forming six convergence clubs

across 81 provinces in Turkey between 2007 and 2019. Regarding the mechanisms through which spatial dependence influences the club clustering process, Ursavaş and Mendez (2022) further demonstrates that the likelihood of club membership is primarily driven by the economic performance of neighboring provinces through spatial externalities in physical capital and structural change. These findings highlight the complex and context-specific relationship between spatial dependence and regional income convergence. Therefore, further research is needed to capture the mechanisms underlying this relationship.

### 3.3 Provincial income disparities and convergence in Vietnam

Numerous studies have addressed the issue of growing income disparities among Vietnamese provinces, including works by Thu Le and Booth (2014); Diez (2016); Bentzen and Tung (2021), and Diem and Hart (2022). Urban–rural income gaps are highlighted as the primary source of income disparities among Vietnamese provinces (Thu Le and Booth 2014; Bentzen and Tung 2021). Concerning income convergence, findings are inconclusive among relevant studies. Hoang (2009) identifies no evidence of absolute  $\beta$ -convergence among Vietnamese provinces during the 1990–2006 period, whereas Minh and Khanh (2013) found weak signs of income convergence between 1991 and 2007 using the Markov chain model. A recent study by Bentzen and Tung (2021) examined the income convergence hypothesis among Vietnamese provinces between 1990 and 2018, utilizing both cross-sectional and time-series approaches. Their findings revealed that approximately half of the Vietnamese provinces were heading to the income standard of Ho Chi Minh City, the leading economy in Southern Vietnam. Using the log  $t$  convergence test, Vu et al (2018) rejected global convergence and identified five convergence clubs among Vietnamese provinces during the 1990–2011 period.

Regarding factors affecting local income clustering, previous studies have shown that fiscal decentralization (Diem and Hart 2022), domestic investment, budget transfers from the central to local government, and foreign direct investment (Vu et al 2018) are among the key determinants of income convergence in Vietnam. A few studies have discussed the effect of spatial dependence on regional income development in Vietnam. Applying global spatial autocorrelation analysis, Hoang (2009) reveals no evidence of spatial dependence in economic growth among Vietnamese provinces over the 1990–2006 period. By contrast, a more recent paper by Esiyok and Ugur (2018) confirms the existence of spatial dependence among Vietnamese provinces' economic growth with the Moran's  $I$  value recorded at 0.2817 for the 2008–2010 period.

Our particular attention is paid to Vu et al (2018). This study utilizes the club convergence framework to investigate the temporal dynamics and determinants of income clustering among Vietnamese provinces. Nevertheless, this study does not tackle the spatial dynamics of local income clustering, creating a research gap for us to fill in.

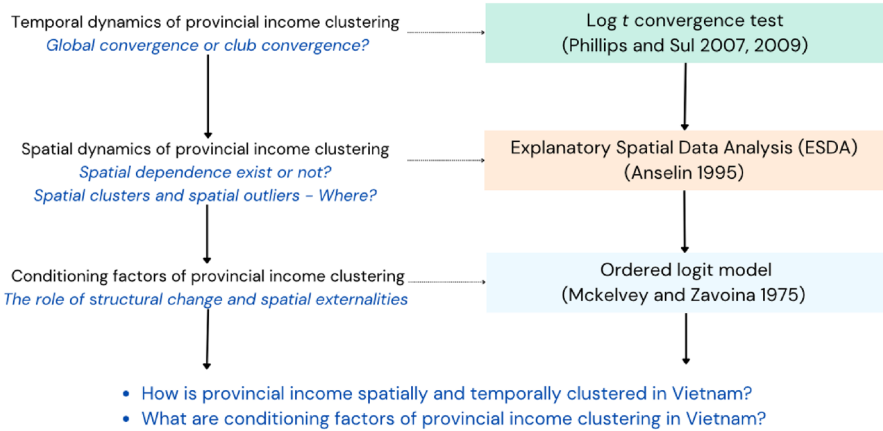


Fig. 1 Study design

## 4 Methodology and data

This paper explores the spatio-temporal dynamics and conditioning factors of regional income convergence in Vietnam, following a three-step research procedure illustrated in Fig. 1.

### 4.1 Log $t$ convergence test and clustering-merging algorithm

#### 4.1.1 Conceptual framework of log $t$ convergence test

The log  $t$  test applied for testing income convergence over a representative set of provincial units starts with the mathematical decomposition of the income variable as follows:

$$y_{it} = \lambda_{it}\mu_t + \varepsilon_{it} \quad (1)$$

where  $y_{it}$  is the logarithm of GDP per capita (ln GDP per capita)<sup>1</sup> of the province  $i$  in year  $t$ ;  $\lambda_{it}$ ,  $\mu_t$ , and  $\varepsilon_{it}$  represent the province-specific time-varying component, the common trend shared by investigated provinces, and the error term, respectively.

Equation (1) can be further transformed into:

<sup>1</sup> It is noted that the logarithm of GDP per capita is abbreviated as "ln GDP per capita" in this study. We apply the analogous abbreviation rule to reflect the logarithmic values of other variables. For instance, "ln Human capital" in Table 2 means the logarithmic value of the human capital variable.

$$y_{it} = \left( \lambda_{it} + \frac{\varepsilon_{it}}{\mu_t} \right) \mu_t = \delta_{it} \mu_t \quad (2)$$

where the province-specific time-varying component and the error term are absorbed in the time-varying idiosyncratic component  $\delta_{it}$ . Equation (2) is a dynamic factor model where the factor-loading coefficient  $\delta_{it}$  represents the distance between the observed variable  $y_{it}$  and the common trend  $\mu_t$ . More intuitively,  $\delta_{it}$  describes the individual transition path of provinces toward the long-run equilibrium  $\mu_t$ .

As it is impossible to fit the model without some restrictions imposed on  $\delta_{it}$  and  $\mu_t$  (Du 2017), Phillips and Sul (2007, 2009) introduce the following semi-parametric specification to characterize the dynamics of  $\delta_{it}$ :

$$\delta_{it} = \delta_i + \frac{\sigma_i \xi_{it}}{\log(t)t^\alpha} \quad (3)$$

where  $\delta_i$  represents the time-independent provincial heterogeneity,  $\sigma_i$  is a province-unique scale parameter,  $\xi_{it}$  is *iid*(0, 1) across  $i$  but may be weakly dependent over  $t$ ,  $\alpha$  is the decay rate. They also define the relative transition parameter,  $h_{it}$ , and the cross-sectional variance,  $H_t$ , as follows:

$$h_{it} = \frac{y_{it}}{N^{-1} \sum_{i=1}^N y_{it}} = \frac{\delta_{it}}{N^{-1} \sum_{i=1}^N \delta_{it}} \quad (4)$$

$$H_t = N^{-1} \sum_{i=1}^N (h_{it} - 1)^2 \quad (5)$$

where  $N$  is the number of investigated provinces. In Eq. (4), by dividing each observed  $y_{it}$  by the panel average, the relative transition parameter eliminates the common component  $\mu_t$  from Eq. (2) (Mendez and Kataoka 2021).

Convergence occurs when all investigated provinces move to the same transition path. Mathematically, the following conditions must be satisfied to achieve convergence among provinces.

$$\begin{aligned} \delta_{it} &\rightarrow \delta \text{ for all } i \text{ as } t \rightarrow \infty \\ h_{it} &\rightarrow 1 \text{ for all } i \text{ as } t \rightarrow \infty \\ H_t &\rightarrow 0 \text{ as } t \rightarrow \infty \end{aligned} \quad (6)$$

Phillips and Sul (2007, 2009) propose to test the null hypothesis of convergence using the log  $t$  regression model as follows:



$$\log\left(\frac{H_1}{H_t}\right) - 2\log(\log(t)) = \hat{a} + \hat{b}\log(t) + \hat{u}_t, \quad (7)$$

for  $t = [rT], [rT] + 1, \dots, T$  and  $r > 0$

where  $r$  is the truncation parameter such that  $rT$  initial observations are discarded and only the  $(1 - r)$  fraction of total observations is kept for regression. As a result of Monte Carlo simulations,  $r$  is suggested to be equal to 0.3 when  $T \leq 50$  (Phillips and Sul 2007, p. 1789). The log  $t$  test is the one-sided heteroskedasticity and autocorrelation consistent  $t$ -test with the null hypothesis of convergence being rejected if the  $t$ -statistic of the  $\hat{b}$  coefficient ( $t_{\hat{b}}$ ) is smaller than  $-1.65$ .

#### 4.1.2 Clustering and merging algorithm

Phillips and Sul (2007, 2009) introduces the clustering and merging algorithm to detect local convergence among subgroups when the null hypothesis of convergence is rejected for the entire sample. A detailed procedure description can be found in Du (2017), besides the original work of Phillips and Sul (2007, 2009). We briefly summarize the procedure as follows:

1. *Sorting*: Sort provinces in descending order according to their observed income values in the last period.
2. *Constructing the core group*: Run the log  $t$  test sequentially for the first  $k$  units ( $2 \leq k < N$ ) and obtain the convergence test statistic  $t_{\hat{b}}(k)$ . The core group is constructed with the group size  $k^*$  decided following the condition:  $k^* = \arg \max_k \{t_{\hat{b}}(k)\}$  subject to  $\min\{t_{\hat{b}}(k)\} > -1.65$ .
3. *Sieving units for club membership*: Add the remaining units to the core group, one at each time, and run the log  $t$  test. The first convergence club, containing the core group members and the ones added, is formulated when the obtained  $t$ -statistic is larger than  $-1.65$ .
4. *Recurring and stopping*: Execute the log  $t$  test over the remaining units that could not be sieved in the previous step. If the obtained  $t$ -statistic is greater than  $-1.65$ , they constitute the second convergence club. Otherwise, rerun steps 1 to 3 to determine any smaller convergence clubs. If no further convergence club is detected, we conclude the remaining units diverge.
5. *Club merging*: Conduct the log  $t$  test for all pairs of subsequent initial clubs. Merge those clubs that jointly fulfill the convergence hypothesis.

#### 4.2 Explanatory Spatial Data Analysis

The Explanatory Spatial Data Analysis (ESDA) method (Anselin 1995) is applied to analyze the spatial dynamics of provincial income clustering in Vietnam. The first ESDA component is global spatial autocorrelation analysis evaluating whether income distribution throughout Vietnam displays a spatially integrated pattern. To

measure global spatial autocorrelation in terms of provincial income distribution, we calculate the global Moran's I statistic in year  $t$  as follows:

$$I_t = \frac{N}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}} \left[ \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (y_{it} - \bar{y}_t) (y_{jt} - \bar{y}_t)}{\sum_{i=1}^N (y_{it} - \bar{y}_t)^2} \right] \quad (8)$$

where  $y_{it}$  and  $y_{jt}$  represent ln GDP per capita of province  $i$  and province  $j$  in year  $t$ , respectively.  $\bar{y}_t$  is the sample average ln GDP per capita in year  $t$ .  $w_{ij}$  is the spatial weight representing the spatial relationship between province  $i$  and province  $j$ . Our study uses the Queen contiguity scheme, one common approach to identifying neighborhood relationships in spatial analysis. This scheme defines two provinces sharing borders or points as neighbors. If province  $i$  and province  $j$  are neighbors,  $w_{ij}$  is equal to 1, and 0 otherwise.  $N$  is the total number of spatial units equal to 63 in this study. The value of the global Moran's I statistic ranges from negative unity to positive unity. The null hypothesis of random spatial distribution is rejected if the global Moran's I statistic is statistically different from 0. A positive (negative) Moran's I statistic indicates a positive (negative) spatial autocorrelation among provinces. That is, nearby provinces display similar (dissimilar) income values.

The second ESDA component is local spatial autocorrelation analysis examining patterns of spatial autocorrelation at specific locations. We calculate the local Moran's I statistic (also known as the Local Indicator of Spatial Association) at province  $i$  in year  $t$  as follows:

$$I_{it} = \frac{(y_{it} - \bar{y}_{it})}{N^{-1} \sum_{i=1}^N (y_{it} - \bar{y}_{it})^2} \sum_{j=1}^N w_{ij} (y_{jt} - \bar{y}_{it}) \quad (9)$$

where the notation is similar to that of Eq. (8). The local spatial autocorrelation analysis indicates locations of "spatial clusters" (i.e., province groups of positive spatial autocorrelation) and "spatial outliers" (i.e., province groups of negative spatial autocorrelation). In the ESDA literature, spatial clusters are further differentiated between the high-high and low-low clusters. While the high-high cluster indicates high-income provinces neighbored by other high-income partners, the low-low cluster indicates low-income provinces surrounded by low-income neighbors. Spatial outliers are also discriminated between the high-low group and the low-high group. The high-low group includes high-income provinces neighbored by low-income partners. In contrast, the low-high group refers to low-income provinces surrounded by high-income neighbors.

### 4.3 Ordered logit regression

We utilize the ordered logit regression to examine conditioning factors of convergence club membership. The variable to explain, denoted by  $c$ , is the club

membership attached to each province. Values of  $c$  are identified according to the clustering-merging algorithm.  $c$  is an ordinal variable because clubs are ranked. Assuming that the club membership  $c$  of province  $i$  is determined by its latent and continuous steady-state income level,  $y_i^*$ , the model explaining conditioning factors can be presented as follows.

$$y_i^* = X_i\beta + \epsilon_i \quad (10)$$

where  $X_i$  represents explanatory variables reflecting provincial characteristics; and  $\epsilon_i$  is the error term having logistic distribution. Our various explanatory variables are organized into four groups. Specifically, based on the standard growth theory, proxy measures of human capital and physical capital are included in the first group, named "growth fundamentals". Population density is included in the second group, named "agglomeration effect". Output shares and changes in output shares of industry and service sectors are included in the third group, named "structural characteristics". Finally, to control for neighboring impacts on provinces' club membership, neighbors' characteristics in terms of GDP per capita, growth fundamentals, agglomeration effect, and sectoral output structures are included in the fourth group, named "spatial externalities". Detailed descriptive statistics of these variables are presented in the subsequent section of the data description. Because the dependent variable  $y_i^*$  is unobserved, we use maximum likelihood estimation to calculate the probabilities of observing values of  $c$  given  $X$ . To quantify the effect of a single factor on the probability of a province joining a particular convergence club, we report the marginal effect on the probabilities of each factor. That is the effect of one unit change in the mean of a single factor on the predicted likelihood of club membership, holding the mean of others constant.

#### 4.4 Data description

This study consists of three components. First, we analyze the temporal dynamics of provincial income clustering in Vietnam using the log  $t$  test. This analysis requires data on the annual real GDP per capita of 63 Vietnamese provinces during the 2010–2020 period. Our study reports provincial GDP per capita in constant 2010 prices. As a prerequisite of the log  $t$  test, we first remove random shocks from the per capita income time series to reveal the trend ln GDP per capita, using the Hodrick-Prescott (HP) filter technique (Hodrick and Prescott 1997). Following the previous literature (Von Lyncker and Thoennessen 2017; Gunawan et al 2021), the smoothing parameter is set equal to 6.25. The descriptive statistics of ln GDP per capita and the trend ln GDP per capita are presented in Table 1.

Next, we examine the spatial dynamics of provincial income clustering in Vietnam using the ESDA methods. This analysis requires calculating the average ln GDP per capita of surrounding neighbors for each of the 63 Vietnamese provinces. Finally, we examine the conditioning factors of income convergence clubs using the ordered logit regression. Data descriptive statistics of club conditioning factors are

**Table 1** Descriptive statistics of provincial ln GDP per capita and the long-run trend ln GDP per capita

Variable	Obs	Mean	SD	Min	Max
ln GDP per capita 2010	63	16.84	0.5	16.18	19.14
Trend ln GDP per capita 2010	63	16.84	0.5	16.19	19.14
ln GDP per capita 2020	63	17.45	0.469	16.61	19.11
Trend ln GDP per capita 2020	63	17.47	0.472	16.62	19.12

**Table 2** Descriptive statistics of club conditioning factors

Variable	Obs	Mean	SD	Min	Max
<i>Growth fundamentals</i>					
ln Human capital	63	2.497	0.39	1.74	3.456
ln Physical capital	63	16.09	0.487	15.12	17.43
<i>Agglomeration effect</i>					
ln Population density	63	5.681	1.012	3.788	8.241
<i>Structural characteristics</i>					
ln Industry share	63	- 1.446	0.451	- 2.376	- 0.185
ln Service share	63	- 0.979	0.307	- 2.338	- 0.429
ln Change in industry share	63	0.317	0.242	- 0.156	1.213
ln Change in service share	63	- 0.049	0.161	- 0.514	0.384
<i>Spatial externalities</i>					
W ln GDP per capita	63	16.83	0.294	16.34	17.65
W ln Human capital	63	2.499	0.284	1.883	2.922
W ln Physical capital	63	16.07	0.23	15.61	16.54
W ln Industry share	63	- 1.458	0.278	- 2.025	- 0.933
W ln Service share	63	- 0.966	0.165	- 1.448	- 0.65
W ln Change in industry share	63	0.317	0.12	0.166	0.747
W ln Change in service share	63	- 0.051	0.099	- 0.285	0.122
W ln Population density	63	5.695	0.765	4.079	7.135

shown in Table 2. Meanwhile, detailed descriptions of the explanatory variables and data sources can be found in Table 11.

Following prior research such as Bartkowska and Riedl (2012) and Ursavaş and Mendez (2022), we have adopted the convention of using the prefix "W" to denote "spatial lags" in our data presentation. This notation is consistently used in Table 2 showing descriptive statistics, Table 11 presenting data description, and other relevant tables. In the field of spatial analysis, the term "spatial lags" refers to geographical neighbors' mean values of specific attributes (Elhorst 2014). In our research, this concept is pivotal for assessing the impacts of spatial externalities on the likelihood of a province joining a specific convergence club. To this end, we define neighboring provinces according to the Queen contiguity scheme and compute neighbors' average values of investigated attributes. In alignment with established practices in previous studies (Bartkowska and Riedl 2012; Cutrini 2019;

**Table 3** Global convergence test result

$\hat{b}$	$t$ -statistic of $\hat{b}$
- 0.87	- 25.63

Ursavaş and Mendez (2022), we record all explanatory variables for the ordered logit regression analysis as they were at the beginning of the observation period, specifically in the year 2010.

## 5 Results and discussions

### 5.1 Temporal dynamics of provincial income clustering

This subsection provides insights into the temporal dynamics of provincial income clustering in Vietnam by discussing the findings obtained from the log  $t$  test. Following previous literature such as Von Lyncker and Thoennessen (2017); Zhang et al (2019), and Ursavaş and Mendez (2022), we first run the log  $t$  test over the whole sample of 63 Vietnamese provinces, i.e.,  $N = 63$ , to check if global income convergence exists. Table 3 presents the result of the log  $t$  test against the null hypothesis of global convergence, with the  $\hat{b}$  coefficient being given in Eq. (7). As can be seen, the null hypothesis of global convergence is rejected at the 95% confidence level ( $t$ -statistic of  $\hat{b} < -1.65$ ), which means that 63 Vietnamese provinces did not converge toward a universal income equilibrium during the 2010–2020 period. Our finding echoes the result reported by Vu et al (2018), which also shows rejection toward the existence of global convergence among Vietnamese localities during the 1990–2011 period.

Since the null hypothesis of global income convergence is rejected, we identify local convergence among province subgroups following the 5-step clustering and merging algorithm. We find local convergence within 7 subgroups with  $t$ -statistic of the  $\hat{b}$  coefficient greater than  $-1.65$  (Table 4). Compared to Vu et al (2018), which reveals 5 income convergence clubs among Vietnamese localities during the 1990–2011 period, our finding of 7 convergence clubs might imply the worsening situation of income discrepancies across Vietnamese provinces. We also identify two provinces showing the diverging trend, namely Dien Bien and Ha Giang. Among possible causes for their income divergence, the most significant might be their lack of a robust economic structure. Situated in the Northern mountainous part of Vietnam, characterized by steep, rugged, and strongly dissected topography, both provinces possess fragile economic structures dominated by small and fragmented agricultural production.

To highlight the discrepancy among clubs, we illustrate the evolution of the average relative trend ln GDP per capita by clubs in Fig. 2. We first remove two diverging provinces, Dien Bien and Lai Chau, to formulate the new sample of convergent provinces. The new sample size ( $N_n$ ) equals 61. Next, we calculate the relative trend value of ln GDP per capita of convergent province  $i$  at time  $t$ , denoted

**Table 4** Identification of final convergence clubs and diverging provinces

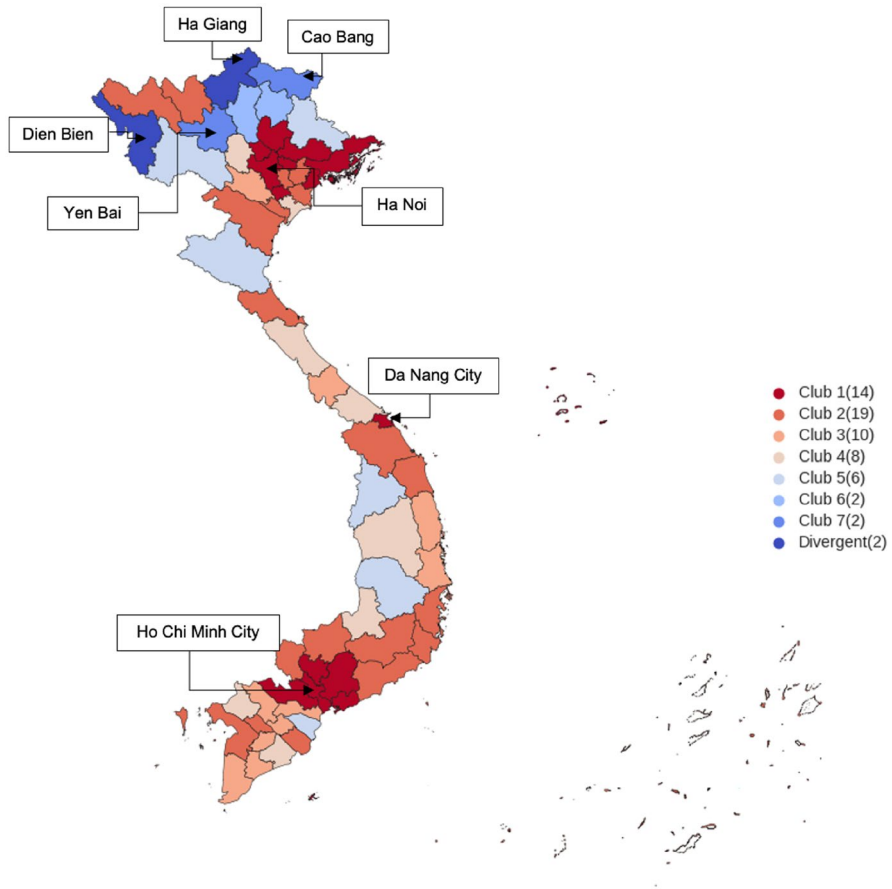
Subgroup [Province number]	Region [Province number]	Membership	$\hat{b}$	<i>t</i> -statistic of $\hat{b}$
Club 1 [14]	Northern Vietnam [7] Central Vietnam [1] Southern Vietnam [6]	Bac Giang, Bac Ninh, Ha Nam Ha Noi, Hai Phong, Thai Nguyen, Vinh Phuc Da Nang Ba Ria - Vung Tau, Binh Duong, Dong Nai, Long An, Quang Ninh, Ho Chi Minh City	- 0.11	- 1.18
Club 2 [19]	Northern Vietnam [6] Central Vietnam [8]	Lai Chau, Lao Cai, Ninh Binh, Thai Binh, Hai Duong, Hung Yen Thanh Hoa, Ha Tinh, Quang Nam, Quang Ngai, Ninh Thuan, Lam Dong, Khanh Hoa, Binh Thuan	- 0.09	- 0.9
Club 3 [10]	Southern Vietnam [5] Northern Vietnam [1] Central Vietnam [3]	Binh Phuoc, Can Tho, Kien Giang, Tay Ninh, Tra Vinh Hoa Binh Binh Dinh, Phu Yen, Quang Tri	0.41	1.86
Club 4 [8]	Southern Vietnam [6] Northern Vietnam [2] Central Vietnam [4] Southern Vietnam [2]	Bac Lieu, Ca Mau, Dong Thap, Hau Giang, Tien Giang, Vinh Long Nam Dinh, Phu Tho Dak Nong, Gia Lai, Quang Binh, Thua Thien - Hue An Giang, Soc Trang	0.24	1.77
Club 5 [6]	Northern Vietnam [3] Southern Vietnam [1]	Lang Son, Son La Dak Lak, Nghe An, Kon Tum Ben Tre	- 0.61	- 1.44
Club 6 [2]	Northern Vietnam [2]	Bac Kan, Tuyen Quang	- 5.11	- 1.41
Club 7 [2]	Northern Vietnam [2]	Cao Bang, Yen Bai	- 0.003	- 0.02
Diverging provinces [2]	Northern Vietnam [2]	Dien Bien, Ha Giang	- 0.38	- 2.55



**Fig. 2** Evolution of the average relative trend in GDP per capita by clubs. *Notes:* The number of club members is shown in round brackets

as  $r(y_{it}) = \frac{t(y_{it})}{N_n^{-1} \sum_{i=1}^{N_n} t(y_{it})}$  where  $t(y_{it})$  is the corresponding trend value, obtained from the HP filter using the smoothing parameter equal 6.25. Using the total province members in each club, the average relative in GDP per capita by clubs is also derived. As can be seen, there are evident discrepancies among the clubs' average relative trend in GDP per capita. The trend line representing Club 1's average income is consistently above the cross-sectional mean (shown by the dashed line). The trend line representing Club 2's average income gradually converges toward the cross-sectional mean, although it remains below the cross-sectional mean line throughout most of the studied period. The other clubs' trend lines of the average relative in GDP per capita are far below the cross-sectional mean. Hence, Club 1 and Club 2 can be categorized as the "high-income" and "middle-income" groups, respectively, while the remaining clubs can be broadly grouped as the "low-income" group. At the end of the target period, the discrepancies between the "high-income" and the "low-income" group seem to increase, confirming the worsening income inequality across Vietnamese provinces during the 2020–2020 period (see Table 9 for details on clubs' average income in 2010 and 2020).

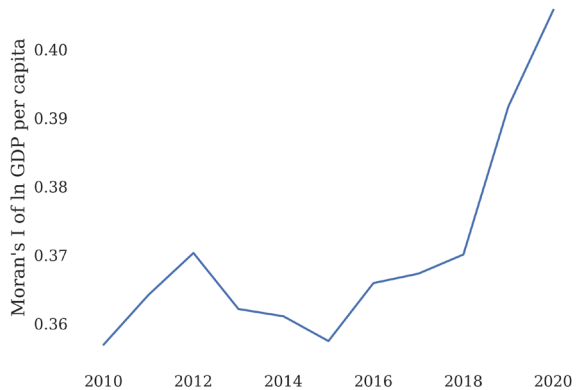
Figure 3 provides the geographical view of club distribution, from which we can elicit three further insights into local income clustering in Vietnam. First, the high-income Club 1 acknowledges the predominant membership of provinces in the Northern KER and Southern KER, whose economic structure is dominated by the industrial sector (see Fig. 7 for a detailed presentation of clubs' sectoral structure). Second, unlike Southern Vietnam, which shows a high degree of economic development homogeneity, Northern Vietnam displays an apparent pattern of economic development heterogeneity among localities. Specifically,



**Fig. 3** Spatial distribution of convergence clubs. *Notes:* The number of club members and diverging provinces is shown in round brackets

while most provinces in the Northern Plain gain membership in the high-income Club 1 and the medium-income Club 2, provinces in the Northern Midlands and Mountains, such as Cao Bang, Yen Bai, Dien Bien, and Ha Giang, are among the country's poorest localities. The homogeneity of economic development in the South and heterogeneity in the North could be partly attributed to the homogeneous and heterogeneous geographical features of Northern and Southern Vietnam, respectively. Finally, nearby provinces tend to have similar income levels. This tendency is particularly noticeable among the "high-income" provinces in Northern and Southern Vietnam and the "low-income" provinces in Northern Vietnam. This motivates us to test the spatial autocorrelation among provinces in the next part of the study.



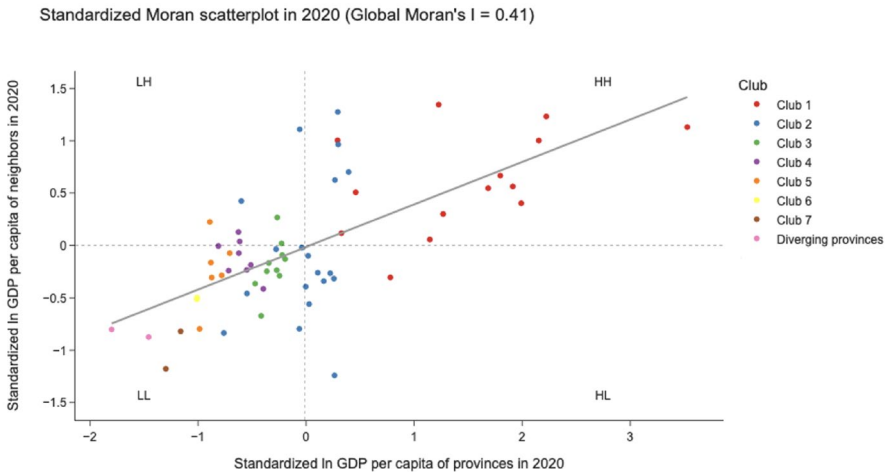
**Fig. 4** Evolution of global spatial dependence

## 5.2 Spatial dynamics of provincial income clustering

In this subsection, we interpret the ESDA results to reveal the spatial dynamics of the local income clustering in Vietnam. It is noted that while interpreting the ESDA results, we employ the terms "spatial dependence", "spatial autocorrelation", and "spatial association" interchangeably.

We first calculate the global Moran's I following Eq. (8) to understand the overall spatial relationship among Vietnamese provinces regarding income distribution. Figure 4 illustrates the direction, magnitude, and evolution of Moran's I statistic between 2010 and 2020. The recorded value of Moran's I statistic increases from 0.36 in 2010 to 0.41 in 2020 and is statistically significant at the 99% confidence level through all years. The statistically significant and positive Moran's I statistic informs us of positive global spatial dependence. In other words, Vietnamese provinces were spatially autocorrelated regarding income distribution during the 2010–2020 period. Adjacent provinces tend to share similar income levels and form clusters in space.

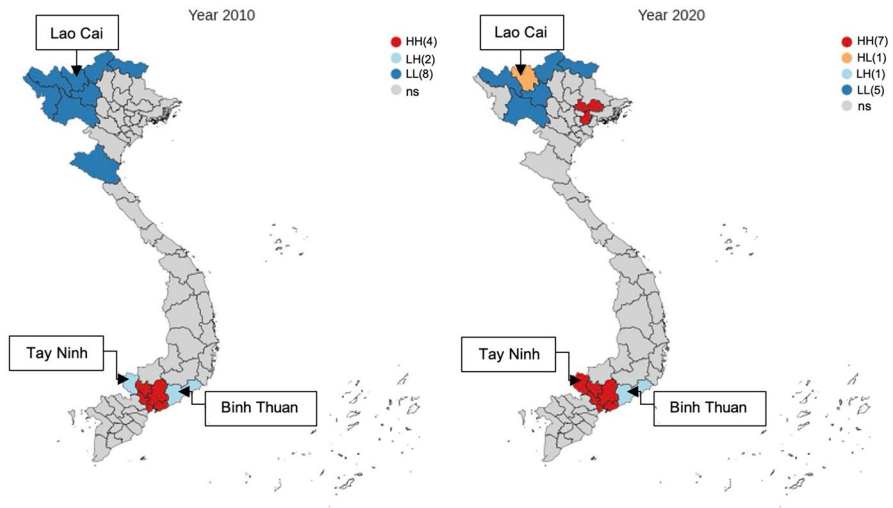
A noteworthy feature is that Moran's I statistic exhibits a significant upward trend between 2010 and 2020. The increase is particularly prominent compared to the previous period when the Moran's I statistic was recorded at 0.2817 (Esiyok and Ugur 2018) between 2008 and 2010. The rising Moran's I statistic signifies the increasing spatial autocorrelation among Vietnamese provinces. This increasing trend can be attributed mainly to strengthened production linkages and enhanced labor mobility, driven by improved regional connectivity in Vietnam during the 2010–2020 period. Specifically, the 2010–2020 decade witnessed the dynamic emergence of KERs as vibrant hubs of foreign direct investment and industrial production following the accomplishment of the KER legal frameworks in the previous decade (refer to Sect. 2 for details). The dynamic industrial activities in these KERs led to a surge in demand for labor and other production inputs. To support those demands, the Vietnamese Government recognized the need to facilitate labor mobility and supply chain integration across provinces by improving regional connectivity. In response to this purpose, the Government made substantial investments in the inter-provincial transportation infrastructure system. Indeed,



**Fig. 5** Global spatial dependence in 2020

improved supply chain integration and labor mobility, which enabled economic synchronization and co-evolution among regions, were recognized as crucial drivers for regional growth spillovers in Vietnam since the commencement of economic reforms (Phan and Coxhead 2010).

In Fig. 5, we plot the club membership to the standardized Moran scatterplot of global spatial dependence in 2020 for further examination of the spatial autocorrelation among club members. Data of provinces' and neighbors' In GDP per capita in 2020 is standardized, i.e., transformed to have a mean of 0 and a standard deviation of 1. Points represent Vietnamese provinces. The X-axis describes the standardized value of provinces' income while the Y-axis represents the standardized value of neighbors' income in 2020. The Moran scatterplot is divided into four quadrants: high-high (HH), low-low (LL), high-low (HL), and low-high (LH). The HH and LL quadrants present spatial grouping among provinces of similar income values. Specifically, the HH (LL) quadrant includes high (low)-income provinces whose geographical neighbors also have high (low) income. The HL and LH quadrants describe the spatial grouping among provinces of dissimilar income values. Specifically, the HL (LH) quadrant presents high (low)-income provinces situated near low (high)-income neighbors. Given positive global spatial autocorrelation, the overall trend line goes through the HH and LL quadrants with a slope equal to 0.41, which is the value of the global Moran's I statistic in 2020. As seen from Fig. 5, almost all members of the high-income group (Club 1) are clustered in the HH quadrant, whereas members of the low-income group (Club 3 to Club 7) are mainly situated in the LL quadrant. Up to this point, we confirm spatial autocorrelation among neighboring provinces, which promotes their transition toward a similar income equilibrium. We will quantify the marginal impact of spatial autocorrelation on the formation of convergence clubs in the next part of analyzing club determinants using the ordered logit model.



**Fig. 6** Local spatial dependence. *Notes:* In this map, HH (LL) represents high- (low-) income provinces bordered by other high- (low-) income neighbors. LH (HL) represents low- (high-) income provinces bordered by other high- (low-) income neighbors. The symbol "ns" denotes provinces with non-significant local spatial dependence. The number of provinces with significant local spatial independence in each quadrant is shown in round brackets

Next, we examine *local spatial autocorrelation* among provinces. Figure 6 displays the areas of statistically significant Local Indicators of Spatial Association (LISA).

First, we concentrate on the low-income cluster, i.e., "Coldspots", in the Northern Midlands and Mountains. Specifically, we delve into Lao Cai Province's development trajectory. Once considered an underprivileged province surrounded by poor neighbors in 2010, Lao Cai witnessed a remarkable income increase by 2020. The substantial income improvement in Lao Cai can be ascribed to multiple factors, notably its distinction as the site for the Lao Cai Border-Gate Economic Zone, one of the most prominent border-gate economic zones in Vietnam. Furthermore, the inception of an expressway linking Lao Cai City directly to Ha Noi has immensely boosted its economic vitality. Capitalizing on these advantageous features, Lao Cai has ascended as a crucial nexus in the Northern mountainous region over the past decade. This presents an opportunity for the Government to mitigate the challenges encountered by the "Coldspots" cluster. Specifically, the Government can foster regional collaborations and economic spillovers within the Northern midlands and mountainous regions by improving infrastructure connectivity between Lao Cai and its contiguous provinces.

Second, we focus on the high-income cluster, i.e., "Hotspots" in Southern Vietnam. Provinces from the Southern KER primarily constitute this cluster. There seem to be economic spillovers among the neighboring provinces within this region. For instance, despite being below the national income average in 2010, Tay Ninh Province ascended to the high-income tier by 2020. This economic leap can be partly attributed to the improved connectivity between Tay Ninh and

**Table 5** Marginal effects of the ordered logit model on club memberships (by industry sector share)

Variables	Club 1	Club 2	Club 3	Club 4	Club 5, 6 and 7
<i>Growth fundamentals</i>					
In Human capital	- 0.01 (0.06)	- 0.02 (0.12)	0.02 (0.09)	0.01 (0.06)	0.01 (0.03)
In physical capital	0.20** (0.10)	0.38** (0.16)	- 0.30** (0.14)	- 0.17* (0.09)	- 0.10* (0.06)
<i>Agglomeration effect</i>					
In Population density	0.09*** (0.04)	0.17*** (0.06)	- 0.14** (0.06)	- 0.08** (0.03)	- 0.05* (0.03)
<i>Structural characteristics</i>					
In industry share	0.20** (0.09)	0.37** (0.15)	- 0.29** (0.14)	- 0.17** (0.08)	- 0.10* (0.06)
In change in industry share	0.34*** (0.12)	0.64*** (0.22)	- 0.51** (0.21)	- 0.30** (0.12)	- 0.17** (0.09)
Observations	61	61	61	61	61

**Notes:** In parentheses are standard errors. Marginal effects are calculated at the means of all explanatory variables. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

other provinces in the Southern KER. Specifically, possessing an extensive border of 240 kms with Cambodia and two critical border-gate economic zones, namely the Moc Bai Border-Gate Economic Zone and the Xa Mat Border-Gate Economic Zone, Tay Ninh serves as an essential conduit for international trade between the Southern KER and other countries in the Mekong Sub-region. Acknowledging the province's strategic significance, the Government channeled considerable resources to strengthen regional connectivity and economic integration within the Southern KER. This is especially evident in two national transportation projects: the expressway linking Tay Ninh with Binh Duong, a leading industrial hub in the South, and the expressway connecting Ho Chi Minh City with the Moc Bai Border-Gate Economic Zone.

Lastly, our focus shifts to Binh Thuan, a low-income province in 2010 located near the Southern KER. Unlike Tay Ninh, Binh Thuan did not advance to the high-income tier by 2020. The two provinces are distinguished by Binh Thuan's non-affiliation with the Southern KER. Probably due to constrained public resources, the Government could not provide sufficient transportation infrastructure for inter-provincial connectivity between the Southern KER and external provinces. This shortfall seemingly impedes the effective economic diffusion from the Southern KER to nearby territories. Our observations suggest that economic spillovers among local economies are intrinsically tied to the quality of regional connectivity. Therefore, implementing a robust transportation framework that enhances inter-provincial linkages is expected to be highly essential for Vietnam's future regional growth.

**Table 6** Marginal effects of the ordered logit model on club memberships (by service sector share)

Variables	Club 1	Club 2	Club 3	Club 4	Club 5, 6 and 7
<i>Growth fundamentals</i>					
In human capital	0.02 (0.05)	0.05 (0.12)	- 0.04 (0.10)	- 0.02 (0.04)	- 0.01 (0.02)
In physical capital	0.19** (0.09)	0.49*** (0.16)	- 0.42*** (0.15)	- 0.17** (0.09)	- 0.08* (0.05)
<i>Agglomeration effect</i>					
In population density	0.08** (0.03)	0.19*** (0.07)	- 0.17** (0.06)	- 0.07** (0.03)	- 0.03* (0.02)
<i>Structural characteristics</i>					
In service share	- 0.24** (0.10)	- 0.61*** (0.19)	0.53*** (0.17)	0.22** (0.11)	0.10 (0.07)
In change in service share	- 0.54*** (0.21)	- 1.37*** (0.42)	1.19*** (0.42)	0.49** (0.21)	0.23* (0.13)
Observations	61	61	61	61	61

**Notes:** In parentheses are standard errors. Marginal effects are calculated at the means of all explanatory variables. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

**Table 7** Marginal effects of the ordered logit model on club memberships (with spatial lags of ln GDP per capita)

Variables	Club 1	Club 2	Club 3	Club 4	Club 5, 6 and 7
<i>Growth fundamentals</i>					
In human capital	0.01 (0.05)	0.02 (0.12)	- 0.02 (0.10)	- 0.01 (0.05)	- 0.00 (0.02)
In physical capital	0.16** (0.08)	0.39** (0.15)	- 0.33** (0.13)	- 0.15* (0.09)	- 0.08* (0.04)
<i>Agglomeration effect</i>					
In population density	0.06** (0.03)	0.15** (0.06)	- 0.13** (0.06)	- 0.06* (0.03)	- 0.03 (0.02)
<i>Structural characteristics</i>					
In industry share	0.13* (0.07)	0.32** (0.16)	- 0.27** (0.13)	- 0.12* (0.07)	- 0.06 (0.04)
In change in industry share	0.26** (0.12)	0.64*** (0.24)	- 0.53** (0.22)	- 0.25** (0.12)	- 0.13* (0.07)
<i>Spatial externalities</i>					
W ln GDP per capita	0.20** (0.10)	0.49** (0.20)	- 0.40** (0.17)	- 0.19* (0.10)	- 0.10 (0.06)
Observations	61	61	61	61	61

**Notes:** In parentheses are standard errors. Marginal effects are calculated at the means of all explanatory variables. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

**Table 8** Marginal effects of the ordered logit model on club memberships (with spatial lags of explanatory variables)

Variables	Club 1	Club 2	Club 3	Club 4	Club 5, 6 and 7
<i>Growth fundamentals</i>					
In human capital	0.09 (0.10)	0.23 (0.23)	- 0.19 (0.21)	- 0.09 (0.09)	- 0.04 (0.04)
In physical capital	0.17* (0.09)	0.43*** (0.16)	- 0.36** (0.16)	- 0.17 (0.11)	- 0.07* (0.04)
<i>Agglomeration effect</i>					
In population density	0.12** (0.05)	0.30** (0.15)	- 0.25* (0.13)	- 0.12* (0.07)	- 0.05 (0.03)
<i>Structural characteristics</i>					
In industry share	0.18** (0.09)	0.46** (0.19)	- 0.38** (0.18)	- 0.18* (0.11)	- 0.07 (0.04)
In change in industry share	0.35*** (0.12)	0.89*** (0.32)	- 0.74** (0.30)	- 0.35* (0.19)	- 0.14* (0.08)
<i>Spatial externalities</i>					
W In Human capital	- 0.12 (0.13)	- 0.29 (0.32)	0.25 (0.27)	0.12 (0.15)	0.05 (0.05)
W In Physical capital	0.36** (0.15)	0.91** (0.38)	- 0.77** (0.37)	- 0.36* (0.19)	- 0.14 (0.09)
W In Population density	- 0.03 (0.05)	- 0.08 (0.14)	0.07 (0.11)	0.03 (0.06)	0.01 (0.02)
W In Industry share	- 0.27 (0.18)	- 0.70* (0.37)	0.59 (0.38)	0.28* (0.17)	0.11 (0.08)
W In Change in industry share	- 0.42 (0.26)	- 1.06* (0.64)	0.89 (0.59)	0.42 (0.28)	0.17 (0.12)
Observations	61	61	61	61	61

*Notes:* In parentheses are standard errors. Marginal effects are calculated at the means of all explanatory variables. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

### 5.3 Conditioning factors of provincial income clustering

This subsection analyzes the conditioning factors of club memberships using the ordered logit model. In the ordered logit regression, two provinces identified as diverging units by the log  $t$  test are removed from our sample, meaning that the total number of observations is 61 (provinces). Because Club 5, Club 6, and Club 7 each contain very few members, we follow the approach of Bartkowska and Riedl (2012) to pool the members of the three last clubs together. Tables 5, 6, 7 and 8 shows the marginal effects of explanatory variables on the probabilities for each club. The first two tables emphasize the role of structural changes, while the latter two highlight the role of spatial externalities on local income development in Vietnam.

Regarding *growth fundamentals*, increasing physical capital enhances the likelihood of provinces transitioning toward a higher income equilibrium. Specifically,

Table 5 shows that one unit of increase in physical capital raises the probability of provinces joining Club 1 (high-income) and Club 2 (middle-income) by 20% and 38%, respectively, while reducing the likelihood of provinces moving to low-income clubs by 30% for Club 3, 17% for Club 4, and 10% for the pooled club of 5, 6, and 7. In the remaining tables (Tables 6, 7, 8), although the magnitude of the marginal effect of physical capital varies, the coefficient signs are robust regardless of the new explanatory variables added to the model. Our findings echo the results of Vu et al (2018), which also confirm the vital contribution of foreign direct and domestic investment to the income development of Vietnam during the 1990–2011 period. As a developing country striving to industrialize, Vietnam shows a high demand for physical capital. Therefore, promoting physical capital accumulation should be a priority of the Government during the country's future economic development.

Regarding *agglomeration effect*, our results on the marginal effect of population density on club memberships are consistent across all models from Tables 5, 6, 7, 8. Populous provinces are more likely to join the high- and middle-income groups (Club 1 and Club 2, respectively). In contrast, provinces with a low population density probably end up with low-income status. This finding is reasonable because populous localities usually enjoy higher purchasing power and labor pooling effects, which are engines for regional economic growth and income improvement. Our finding is consistent with the results of Cutrini (2019), indicating that metropolitan European regions are generally labeled with a high-income status. The population density in Vietnam is highest among the Northern KER and Southern KER provinces, which dominate Club 1 and Club 2.

To fully understand the impacts of *structural characteristics* on local income clustering in Vietnam, we separately test our models with two processes: transformation toward the industry-oriented structure (Table 5) and shift toward the service-oriented structure (Table 6). Both the initial sector shares and the sectoral change dominantly affect club memberships. However, having larger magnitudes of marginal effects, the sector change seems to outweigh the initial sector share in explaining club memberships. Specifically, Table 5 shows that one unit of increase in the change of the industry share will grow the likelihood of joining Club 1 (high income) by 34% and Club 2 (middle income) by 64%, compared to the respective increases of 20% and 37% resulting from one unit of increase in the initial industry share. This suggests an important policy implication. Industrial acceleration is necessary for the significant development of provincial income in Vietnam. Regardless of initial sectoral composition, provinces may still enjoy income improvement if they can successfully conduct the structural transformation.

It should be noted that different directions of structural transformation will have opposite impacts on Vietnam's local income development. Table 6 indicates that provinces having an initial service-dominant structure or performing a transformation toward the service-dominant structure face a higher probability of ending up in the lower-income clubs. This finding is explainable in the current context of the Vietnamese economy. The service sector in Vietnam is currently

dominated by low-paid service sectors.<sup>2</sup> Studying the sectoral structure of each club, we also find that service dominates the economic structure of the low-income group (Club 3 to Club 7) (see Fig. 7). The results shown in Table 6 imply that an economic structure prevailed by a low-paid service sector decreases the likelihood of provinces obtaining a higher income status. At this point, our findings are consistent with the results of Cutrini (2019) and Ursavaş and Mendez (2022) of the negative impacts of routine services on income upgrading across European regions and Turkish territories, respectively. In the future, moving laborers from low-productivity service sectors to higher-productivity industrial industries is essential for local income improvement in Vietnam.

Finally, we discuss the role of *spatial externalities* on club memberships. In the previous section, the positive global spatial autocorrelation among nearby provinces has been confirmed. Now we quantify how spatial autocorrelation affects club memberships by adding the spatial lag of ln GDP per capita ( $W \ln \text{GDP per capita}$ ) to the baseline model. Results are shown in Table 7. We find statistically significant coefficients of  $W \ln \text{GDP per capita}$  assigned to Club 1, Club 2, Club 3, and Club 4. Specifically, provinces surrounded by high-income neighbors will be more likely to join the high-income Club 1 and the middle-income Club 2 by 20% and 49%, respectively, while less likely to register in the low-income Club 3 by 40% and Club 4 by 19%. Although the coefficient of  $W \ln \text{GDP per capita}$  assigned to the last pooled group is found statistically insignificant, the negative coefficient sign reflects our expectation that a province sharing borders with rich neighbors will be less likely to have a low-income status.

After finding evidence of spatial income spillovers, the next question is through which mechanisms income spatial spillovers occur. To answer this question, we add spatial lags of other explanatory variables reflecting characteristics of neighboring provinces to the baseline model. The results presented in Table 8 indicate that the spatial income spillovers possibly originate from spatial investment spillovers. Specifically, one unit of increase in physical capital accumulation in neighboring provinces will increase the likelihood of focal provinces obtaining memberships in the high-income Club 1 and the middle-income Club 2 by 36% and 91%, respectively. Our finding is supported by the theory line pioneered by Mankiw et al (1992), who prove that regional economies benefit from increased economic activities in nearby regions through capital investment. López-Bazo et al (2004) emphasizes that spillovers from the increased physical capital in nearby regions could compensate for the diminishing return rate to the scale of capital accumulation in prosperous economies and allow them to maintain economic growth. In the spatial economic literature, this “win-win” relationship is defined as “*spatial complementarity*”.

Contrary to the spatial complementarity in terms of physical capital accumulation, our results suggest *spatial competition* among adjacent provinces during industrial intensification. Regarding the coefficients attached to the spatial lags of the initial industry share ( $W \ln \text{Industry share}$ ) and the spatial lag of the change in industry share ( $W \ln \text{Change in industry share}$ ), the results show that having neighbors possessing higher initial industrial shares or intensifying industrial activities could be a detriment to the economic

<sup>2</sup> According to data provided by the General Statistics Office (GSO), high-paid service categories, such as information and media, finance and banking, scientific and technological activities, education and training only accounted for 29.21% of the total service sector in 2020.



development of the focal province. This is possible because contiguous areas might compete to attract skilled workers or foreign direct investment during industrial development. Provinces facing labor-force drainage will be less advantageous in achieving a higher income status. In the context of the Vietnamese economy, previous literature also indicates the competitive relationship among Vietnamese provinces in attracting foreign direct investment (Esiyok and Ugur 2017).

## 6 Conclusions and policy recommendations

In this article, we propose a combined approach of log  $t$  test, Explanatory Spatial Data Analysis (ESDA), and ordered logit model to comprehensively reveal the spatio-temporal dynamics and conditioning factors of provincial income clustering in Vietnam between 2010 and 2020.

We first show that seven convergence clubs characterize provincial income clustering in Vietnam. Increasing gaps among clubs' average trend income signify growing regional income inequalities in Vietnam. Provinces sharing the same club memberships do not scatter across space. Instead, we find a clear pattern of spatial autocorrelation, which encourages nearby provinces to transition toward sharing income equilibria. We also identify two high-income clusters (i.e., "Hotspots"), including provinces of the Northern KER and Southern KER, and one low-income cluster (i.e., "Coldspots") in the Northern mountainous areas. We believe that this act of navigation is highly beneficial, helping the Government detect and tackle persistently poor localities to reduce income disparities in Vietnam.

Regarding conditioning factors of convergence clubs, we reveal physical capital and structural change as two vital internal factors affecting provinces' potential transition paths. It should be noted that only the transformation toward a structure dominated by higher value-added industrial sectors can help provinces upgrade their income level. Besides revealing the essential internal factors, we highlight the outsized role of spatial externalities among provinces as the key external factor in defining club memberships. Specifically, nearby provinces interact and influence each other via spatial complementarity in physical accumulation and spatial competition in industrial intensification. These new findings deepen the current knowledge about factors affecting the local income clustering in Vietnam and allow us to develop practical policy implications for regional income improvement.

We encourage the Vietnamese Government to exploit spatial dependence while tackling regional income inequalities. For example, in addition to promoting physical accumulation and structural change across local economies, the Government may think of focalizing public investment and, at the same time, developing good-quality connectivity infrastructure among the focalized and other surrounding provinces to exaggerate spillover effects among them. Linkages among regional economies are a prerequisite of spatial externalities. Finally, because of the spatial complementarity in physical capital accumulation and spatial competition in industrial intensification among provinces, the Government may think of creating a policy coordination

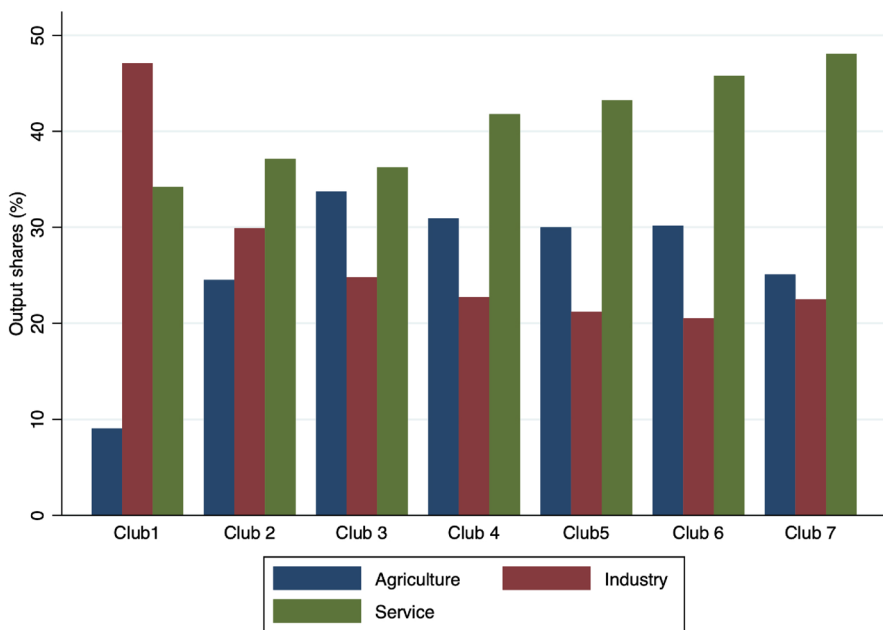
framework to intensify the complementary externalities while reducing potential competition among provinces during their industrialized economic development.

## Appendix

See below Table 9, Fig. 7, Tables 10, 11, Fig. 8 and Table 12.

**Table 9** Club average GDP per capita, 2010–2020

Subgroup [Province number]	Average GDP per capita 2010 (million VND/person/year)	Average GDP per capita 2020 (million VND/person/year)
Club 1 [14]	50.93	82.94
Club 2 [19]	19.52	38.28
Club 3 [10]	18.13	32.92
Club 4 [8]	16.10	28.57
Club 5 [6]	15.41	25.43
Club 6 [2]	14.03	23.61
Club 7 [2]	13.84	21.31
Diverging provinces [2]	11.63	17.72



**Fig. 7** Sectoral output structure of clubs

**Table 10** Correlation matrix

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) ln Human capital	1.000							
(2) ln Physical capital	0.649***	1.000						
(3) ln Population density	0.224*	0.262**	1.000					
(4) ln Industry share	0.428***	0.559***	0.337***	1.000				
(5) ln Service share	0.262**	-0.115	-0.227*	-0.513***	1.000			
(6) ln Change in industry share	-0.252**	-0.307**	-0.124	-0.419***	0.151	1.000		
(7) ln Change in service share	-0.204	0.001	-0.014	-0.142	-0.273**	-0.612***	1.000	
(8) W ln GDP per capita	0.120	0.264**	0.399***	0.408***	-0.444***	-0.131	0.019	1.000

\*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

**Table 11** Variables and data sources

Variable	Explanations	Data source
<i>Growth fundamentals</i>		
In Physical capital	Logarithmic investment per capita in 2010 (million VND/person/year)	GSO, Vietnam
In Human capital	Logarithmic share of labors receiving official training degrees or certificates to total labors in 2010	GSO, Vietnam
<i>Agglomeration effect</i>		
In Population density	Logarithmic number of people per square kilometer in 2010	GPW, v4, Geoquery
<i>Structural changes</i>		
In Industry share	Logarithmic share of industry plus construction sectors to total GDP in 2010	GSO, Vietnam
In Service share	Logarithmic share of service sectors (with knowledge- intensive service sectors included) to total GDP in 2010	GSO, Vietnam
In Change in industry share	Difference in logarithmic shares of industry plus construction sectors between 2010 and 2020	GSO, Vietnam
In Change in service share	Difference in logarithmic shares of service sectors (with knowledge-intensive service sectors included) between 2010 and 2020	GSO, Vietnam
<i>Spatial externalities</i>		
W In GDP per capita	Spatial lags of GDP per capita, that is, logarithmic GDP per capita of neighbors in 2010	Own calculations
W In Physical capital	Spatial lags of physical capital, that is, logarithmic investment per capita of neighbors in 2010	Own calculations
W In Human capital	Spatial lags of human capital, that is, logarithmic share of labors with official training degrees or certificates to total labors of neighbors in 2010	Own calculations
W In Industry share	Spatial lags of industry share, that is, the logarithmic share of industry plus construction sectors to total GDP of neighbors in 2010	Own calculations
W In Service share	Spatial lags of service share, that is, the logarithmic share of service sectors (with knowledge-intensive service sectors included) to total GDP of neighbors in 2010	Own calculations
W In Change in industry share	Spatial lags of change in industry share, that is, the difference in the logarithmic shares of industry plus construction sectors of neighbors between 2010 and 2020	Own calculations
W In Change in service share	Spatial lags of change in service share, that is, the difference in the logarithmic shares of service sectors with knowledge-intensive service sectors included) of neighbors between 2010 and 2020	Own calculations
W In Population density	Spatial lags of population density, that is, the logarithmic number of people per square kilometer of neighbors in 2010	Own calculations

**Notes:** GSO stands for General Statistics Office of Vietnam. GPW, v4 stands for the Gridded Population of the World, Version 4, provided by Center for International Earth Science Information Network, Columbia University

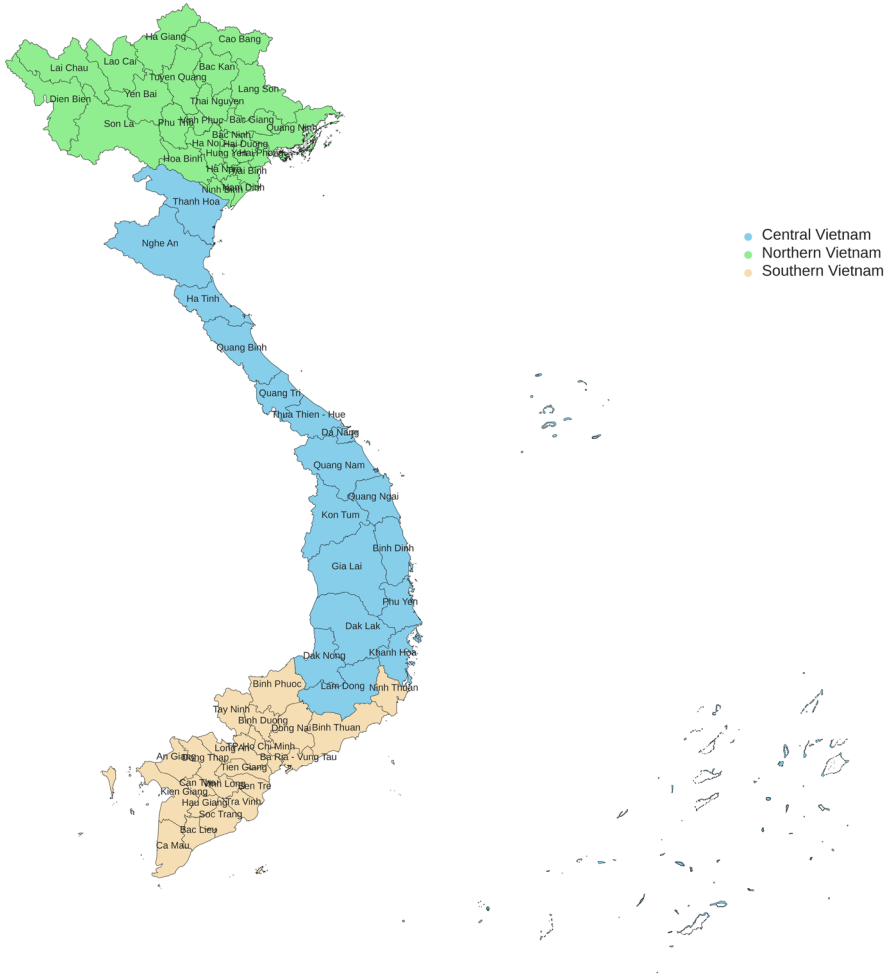


Fig. 8 Three geographical regions of Vietnam

**Table 12** Overview of the main economic features of three KERs

Regions	Main economic features
<i>Northern KER</i>	<p><i>Number of provinces:</i> 8 provinces (Ha Noi, Hung Yen, Hai Phong, Quang Ninh, Hai Duong, Bac Ninh, Vinh Phuc)</p> <p><i>Core municipality:</i> Ha Noi</p> <p><i>Main sectors:</i> automobiles and motorcycle parts, electronics, electricity, shipbuilding, logistics, textiles, supporting industries</p>
<i>Central KER</i>	<p><i>Number of provinces:</i> 5 provinces (Thua Thien-Hue, Da Nang, Quang Nam, Quang Ngai, Binh Dinh)</p> <p><i>Core municipality:</i> Da Nang</p> <p><i>Main sectors:</i> seaports and logistics, seafood and food processing, tourism, shipbuilding and maritime services</p>
<i>Southern KER</i>	<p><i>Number of provinces:</i> 8 provinces (Ho Chi Minh City, Binh Duong, Ba Ria-Vung Tau, Dong Nai, Tay Ninh, Binh Phuoc, Long An, Tien Giang)</p> <p><i>Core municipality:</i> Ho Chi Minh City</p> <p><i>Main sectors:</i> electronics, IT and software, telecommunications, food processing, oil and gas, logistics, chemical and supporting industries</p>

**Acknowledgements** The author thanks Prof. Otsubo Shigeru, Prof. Mendez Carlos, and Prof. Christian Otchia of the Graduate School of International Development, Nagoya University, Japan, for their constant supervision of her research. Heartfelt thanks also go to Prof. Yoshiro Higano, the Editor-in-Chief of the Asia-Pacific Journal of Regional Science, and two anonymous referees for their constructive and insightful comments, which have substantially improved the quality of this article. The author would like to acknowledge, with gratitude, the significant support from members of the QuaRCS Lab Japan while conducting this study. The author is the recipient of the Japanese Government (MEXT) Scholarship for her PhD study (2022-2025).

**Funding** The author did not receive support from any organization for the submitted work. Open Access funding by the Japan Section of Regional Science Association International.

**Data and Code availability** Available upon reasonable request.

## Declarations

**Conflict of interest** The author declares no competing interests that are relevant to the content of this article.

**Ethics approval and Consent** Not applicable.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aginta H (2021) Spatial dynamics of consumer price in Indonesia: convergence clubs and conditioning factors. *Asia Pac J Reg Sci* 5(2):427–451. <https://doi.org/10.1007/s41685-020-00178-0>
- Anselin L (1995) Local indicators of spatial association-LISA. *Geogr Anal* 27(2):93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Bartkowska M, Riedel A (2012) Regional convergence clubs in Europe: Identification and conditioning factors. *Econ Model* 29(1):22–31. <https://doi.org/10.1016/j.econmod.2011.01.013>
- Bentzen J, Tung LT (2021) Regional income convergence in an emerging Asian economy: empirical evidence from Vietnam. *Post Commun Econ* 33(1):64–78. <https://doi.org/10.1080/14631377.2020.1722587>
- Cutrini E (2019) Economic integration, structural change, and uneven development in the European Union. *Str Chang Econ* 50:102–113. <https://doi.org/10.1016/j.strueco.2019.06.007>
- Diem LTT, Hart N (2022) Fiscal decentralization and income convergence: evidence from Vietnam. *J Asia Pac Econ*. <https://doi.org/10.1080/13547860.2022.2108211>
- Diez JR (2016) Vietnam 30 years after Doi Moi: achievements and challenges. *Zeitschrift für Wirtschaftsgeographie* 60(3):121–133. <https://doi.org/10.1515/zfw-2016-0035>
- Du K (2017) Econometric convergence test and club clustering using Stata. *Stata J* 17(4):882–900. <https://doi.org/10.1177/1536867X1801700407>
- Elhorst JP (2014) *Spatial econometrics: from cross-sectional data to spatial panels*. Springer, Heidelberg
- Esiyok B, Ugur M (2017) A spatial regression approach to FDI in Vietnam: Province-level evidence. *Singapore Econ Rev* 62(02):459–481. <https://doi.org/10.1142/S0217590815501155>
- Esiyok B, Ugur M (2018) Spatial dependence in the growth process and implications for convergence rate: evidence on Vietnamese provinces. *J Asia Pac Econ* 23(1):51–65. <https://doi.org/10.1080/13547860.2017.1351764>
- GSO (2020) *Statistical Year Book of Vietnam 2019*. Statistical Publishing House, Hanoi
- Gunawan AB, Mendez C, Otsubo S (2021) Provincial income convergence clubs in Indonesia: Identification and conditioning factors. *Growth Change* 52(4):2540–2575. <https://doi.org/10.1111/grow.12553>
- Hoang NH (2009) *Regional welfare disparities and regional economic growth in Vietnam*. Dissertation. Wageningen University and Research
- Hodrick RJ, Prescott EC (1997) Postwar US business cycles: an empirical investigation. *J Money Credit Bank* 29(1):1–16. <https://doi.org/10.2307/2953682>
- Islam N (2003) What have we learnt from the convergence debate? *J Econ Surv* 17(3):309–362. <https://doi.org/10.1111/1467-6419.00197>
- López-Bazo E, Vayá E, Artis M (2004) Regional externalities and growth: evidence from European regions. *J Reg Sci* 44(1):43–73. <https://doi.org/10.1111/j.1085-9489.2004.00327.x>
- Mankiw NG, Romer D, Weil DN (1992) A contribution to the empirics of economic growth. *Q J Econ* 107(2):407–437. <https://doi.org/10.2307/2118477>
- Martin V, Vazquez G (2015) Club convergence in Latin America. *BE J Macroecon* 15(2):791–820. <https://doi.org/10.1515/bejm-2014-0109>
- Mendez C, Kataoka M (2021) Disparities in regional productivity, capital accumulation, and efficiency across Indonesia: A club convergence approach. *Rev Dev Econ* 25(2):790–809. <https://doi.org/10.1111/rode.12726>
- Minh NK, Khanh P (2013) Forecasting the convergence state of per capita income in Vietnam. *Am J Oper Res* 3(06):487–496. <https://doi.org/10.4236/ajor.2013.36047>
- Oxfam, (2017) *Even it Up: How to tackle inequality in Vietnam*. Labor-Social Publishing House, Hanoi
- Pham TH, Riedel J (2019) Impacts of the sectoral composition of growth on poverty reduction in Vietnam. *J Econ Dev* 21(2):213–222. <https://doi.org/10.1108/JED-10-2019-0046>
- Phan D, Coxhead I (2010) Inter-provincial migration and inequality during Vietnam's transition. *J Dev Econ* 91(1):100–112. <https://doi.org/10.1016/j.jdeveco.2009.06.008>
- Phillips PC, Sul D (2007) Transition modeling and econometric convergence tests. *Econometrica* 75(6):1771–1855. <https://doi.org/10.1111/j.1468-0262.2007.00811.x>
- Phillips PC, Sul D (2009) Economic transition and growth. *J Appl Econ* 24(7):1153–1185. <https://doi.org/10.1002/jae.1080>
- Tam PS (2018) Economic transition and growth dynamics in Asia: harmony or discord? *Comp Econ Stud* 60:361–387. <https://doi.org/10.1057/s41294-018-0051-y>

- Le Thu H, Booth AL (2014) Inequality in Vietnamese urban-rural living standards, 1993–2006. *Rev Income Wealth* 60(4):862–886. <https://doi.org/10.1111/roiw.12051>
- Ursavaş U, Mendez C (2022) Regional income convergence and conditioning factors in Turkey: revisiting the role of spatial dependence and neighbor effects. *Ann Reg Sci*. <https://doi.org/10.1007/s00168-022-01168-0>
- Von Lyncker K, Thoennesen R (2017) Regional club convergence in the EU: evidence from a panel data analysis. *Empir Econ* 52(2):525–553. <https://doi.org/10.1007/s00181-016-1096-2>
- Vu BX, Hoang VNV, Nghiem S (2018) Provincial divergence and sub-group convergence in Vietnam's GDP per capita. *J Econ Res* 23(1):81–107
- Zhang W, Xu W, Wang X (2019) Regional convergence clubs in China: identification and conditioning factors. *Ann Reg Sci* 62(2):327–350. <https://doi.org/10.1007/s00168-019-00898-y>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.